

# DATA PROCESSOR, ITS DATA PROCESSING METHOD, AND STORAGE MEDIUM STORING COMPUTER-READABLE PROGRAM

Publication number: JP11085457

Publication date: 1999-03-30

Inventor: IKEDA HIROAKI

Applicant: CANON KK

Classification:

- international: G06F3/16; G06F3/048; G06F3/14; G06K9/03;  
G06K9/20; G06K9/20; G06F3/16; G06F3/048;  
G06F3/14; G06K9/03; G06K9/20; G06K9/20; (IPC1-7):  
G06F3/16; G06F3/14; G06K9/03; G06K9/20

- European:

Application number: JP19970246351 19970911

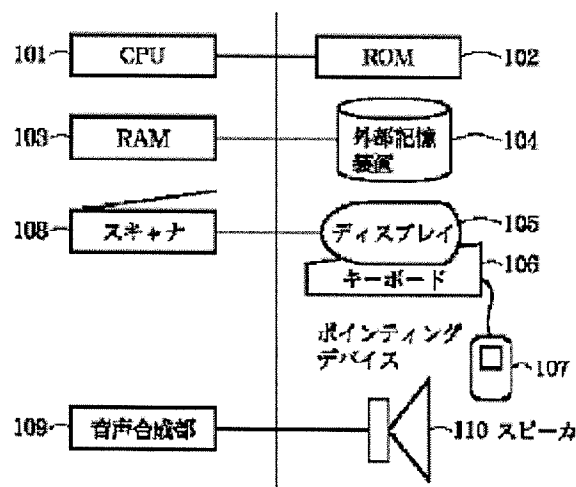
Priority number(s): JP19970246351 19970911

Report a data error here

## Abstract of JP11085457

**PROBLEM TO BE SOLVED:** To enable an operator to quickly confirm the contents of an original image by voice without displaying an image of the whole original image.

**SOLUTION:** A CPU 101 analyzes a read original image, extracts a rectangular area of an optional size, adds each different attribute data to the rectangular area based on the form of each extracted rectangular area, analyzes image data in a rectangular area, to which attribute data are added to recognize character information and stores the character information and the attribute information in an external storage device 104 together with the image data. Then, the CPU 101 specifies an original image to be retrieved and its attributes out of stored original images, judges whether or not an area based on the specified original image and attribute are selected from an external storage device 104 or not and controls the voice reading of character information added to area information which is executed by a voice synthesis part 109 through a speaker 110 based on the judged result.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-85457

(43) 公開日 平成11年(1999) 3月30日

| (51) Int.Cl. <sup>6</sup> | 識別記号  | F I                  |
|---------------------------|-------|----------------------|
| G 0 6 F 3/16              | 3 3 0 | G 0 6 F 3/16 3 3 0 C |
|                           | 3 1 0 | 3/14 3 1 0 A         |
| G 0 6 K 9/03              |       | G 0 6 K 9/03 B       |
| 9/20                      | 3 4 0 | 9/20 3 4 0 L         |

審査請求 未請求 請求項の数12 O L (全 14 頁)

(21) 出願番号 特願平9-246351

(22) 出願日 平成9年(1997) 9月11日

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 池田 裕章

東京都大田区下丸子3丁目30番2号 キヤ  
ノン株式会社内

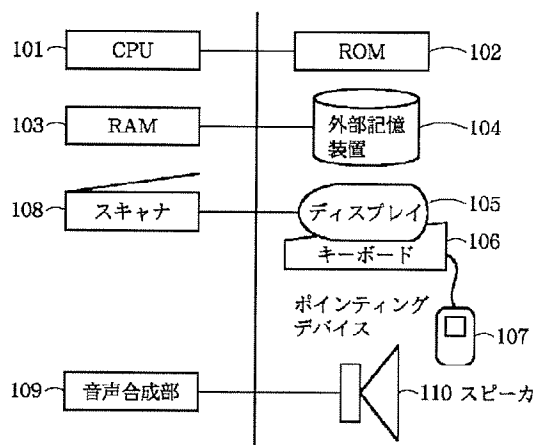
(74) 代理人 弁理士 小林 将高

(54) 【発明の名称】 データ処理装置およびデータ処理装置のデータ処理方法およびコンピュータが読み出し可能なプログラムを格納した記憶媒体

(57) 【要約】

【課題】 原稿画像全体のイメージを表示させなくても、短時間の内に原稿画像の内容を音声でオペレータに確認させること。

【解決手段】 CPU 101が読み取られた原稿画像を解析して任意のサイズの矩形領域に抽出し、該抽出された各矩形領域の形態に基づいて異なる属性データを付加し、該属性データが付加された矩形領域内の画像データを解析して文字情報を認識し、該文字情報と属性データを付加して前記画像データと共に外部記憶装置104に保存し、該保存された原稿画像に対して検索すべき原稿画像および属性を指定し、該指定された原稿画像および属性に基づく領域が外部記憶装置104から選択されているかどうかを判定した結果に基づいて、音声合成部109がスピーカ110を介して行う領域情報に付加されている文字情報の音声読み上げを制御する構成を特徴とする。



## 【特許請求の範囲】

【請求項 1】 原稿上の文書画像を読み取り画像データを出力する読取り手段と、

前記読取り手段により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出手段と、

前記抽出手段により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加手段と、

前記付加手段により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識手段と、

各領域に対して前記認識手段により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存手段と、を有することを特徴とするデータ処理装置。

【請求項 2】 文字情報を音声として出力する出力手段と、

前記メモリ資源に保存された画像データに対して検索すべき領域の属性を指定する第 1 の指定手段と、

前記第 1 の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 1 の判定手段と、

前記第 1 の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第 1 の制御手段と、を有することを特徴とする請求項 1 記載のデータ処理装置。

【請求項 3】 前記抽出手段により抽出された矩形領域および前記認識手段により認識された文字情報を表示部に表示させる第 1 の表示制御手段と、

前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示手段と、

前記指示手段による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識手段により認識された文字情報を修正する編集手段と、を有することを特徴とする請求項 1 記載のデータ処理装置。

【請求項 4】 文字情報を音声として出力する出力手段と、

前記メモリ資源に保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第 2 の表示制御手段と、

前記表示部に表示される原稿画像および属性項目から所望の原稿画像および属性項目を指定する第 2 の指定手段と、

前記第 2 の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 2 の判定手段と、

前記第 2 の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第 2 の制御手段と、を有することを特徴とする請求項 3 記載のデータ処理装置。

【請求項 5】 原稿上の文書画像を読み取り画像データを出力する読取り工程と、

前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程と、

前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程と、

前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程と、

各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程と、を有することを特徴とするデータ処理装置のデータ処理方法。

【請求項 6】 原稿上の文書画像を読み取り画像データを出力する読取り工程と、

前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程と、

前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程と、

前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程と、

各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程と、

を有することを特徴とするコンピュータが読み出し可能なプログラムを格納した記憶媒体。

【請求項 7】 前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第 1 の指定工程と、

前記第 1 の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 1 の判定工程と、

前記第 1 の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程と、を有することを特徴とする請求項 5 記載のデータ処理装置のデータ処理方法。

【請求項 8】 前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第 1 の指定工程と、

前記第 1 の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 1 の判定工程と、

前記第 1 の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程と、を有することを特徴とする請求項 6 記載のコンピュータが読み出し可能なプログラムを格納した記憶媒体。

【請求項 9】 前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第 1 の表示工程と、

前記表示部により表示された矩形領域または文字情報に

に対する修正指示を行う指示工程と、  
前記指示工程による修正指示に基づいて前記抽出工程により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程と、を有することを特徴とする請求項 5 記載のデータ処理装置のデータ処理方法。

【請求項 1 0】 前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第 1 の表示工程と、  
前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示工程と、  
前記指示工程による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程と、を有することを特徴とする請求項 6 記載のコンピュータが読み出し可能なプログラムを格納した記憶媒体。

【請求項 1 1】 前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第 2 の表示工程と、  
前記表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第 2 の指定工程と、  
前記第 2 の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 2 の判定工程と、  
前記第 2 の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程と、  
を有することを特徴とする請求項 9 記載のデータ処理装置のデータ処理方法。

【請求項 1 2】 前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第 2 の表示工程と、  
前記表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第 2 の指定工程と、  
前記第 2 の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 2 の判定工程と、  
前記第 2 の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程と、を有することを特徴とする請求項 1 0 記載のコンピュータが読み出し可能なプログラムを格納した記憶媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】 本発明は、原稿画像を読み取って抽出される領域毎に特定の属性を付加して保存し、該保存された領域の文字情報を音声出力可能なデータ処理装置およびデータ処理装置のデータ処理方法およびコンピュータが読み出し可能なプログラムを格納した記憶媒体に関するものである。

【0 0 0 2】

【従来の技術】 従来の電子ファイリング機能を備えるデ

ータ処理装置では、スキャナ等から入力された原稿画像を記憶装置に記憶させる場合に、ユーザにより指定されたキーワード等を原稿全体と対応付けて記憶装置に記憶させていた。また、記憶された原稿画像を検索する場合には、キーワードを入力して記憶されたキーワードと合致する原稿を記憶装置から検索し、該当する原稿画像があればその原稿画像を読み出して、ディスプレイ等に表示させるように構成されている。

【0 0 0 3】 また、アプリケーションで作成された文書情報をファイリングするデータ処理装置においても、あらかじめ各文書に対応付けて付加したキーワードあるいは文書全文を利用して検索処理するように構成されている。

【0 0 0 4】

【発明が解決しようとする課題】 従来のデータ処理装置は上記のように構成されているので、記憶装置に記憶させた原稿画像の内容を参照する場合に、特にその内容を大まかに把握する場合には、原稿画像の一部の内容が得られればその原稿を把握できるのにもかかわらず、原稿の全ての内容を再現する処理を実行してしまうので内容を確認するまでに多くの時間が必要であるという問題点があった。

【0 0 0 5】 本発明は上記の問題点を解消するためになされたもので、本発明の目的は、読み取られた原稿画像を解析して任意のサイズの矩形領域に抽出し、該抽出された各矩形領域の形態に基づいて異なる属性データを付加し、該属性データが付加された矩形領域内の画像データを解析して文字情報を認識し、該文字情報と属性データを付加して前記画像データと共にメモリ資源に保存し、該保存された原稿画像に対して検索すべき原稿画像および属性を指定し、該指定された原稿画像および属性に基づく領域がメモリ資源から選択されているかどうかを判定した結果に基づいて、領域情報に付加されている文字情報の音声読み上げを制御することにより、原稿画像の領域に付加されている文字情報を音声として読み上げて、原稿画像全体のイメージを表示させるための多大な処理時間を必要とすることなく、オペレータの特定の属性および原稿画像を指定するという簡単な操作だけで、短時間の内に原稿画像の内容を音声でオペレータに確認させることである。

【0 0 0 6】

【課題を解決するための手段】 本発明に係る第 1 の発明は、原稿上の文書画像を読み取り画像データを出力する読取り手段と、前記読取り手段により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出手段と、前記抽出手段により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加手段と、前記付加手段により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識手段と、各領域に対して前記認識手

段により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存手段とを有するものである。

【0007】第2の発明は、文字情報を音声として出力する出力手段と、前記メモリ資源に保存された画像データに対して検索すべき領域の属性を指定する第1の指定手段と、前記第1の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定手段と、前記第1の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第1の制御手段とを有するものである。

【0008】第3の発明は、前記抽出手段により抽出された矩形領域および前記認識手段により認識された文字情報を表示部に表示させる第1の表示制御手段と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示手段と、前記指示手段による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識手段により認識された文字情報を修正する編集手段とを有するものである。

【0009】第4の発明は、文字情報を音声として出力する出力手段と、前記メモリ資源に保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示制御手段と、前記表示部に表示される原稿画像および属性項目から所望の原稿画像および属性項目を指定する第2の指定手段と、前記第2の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定手段と、前記第2の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第2の制御手段とを有するものである。

【0010】第5の発明は、原稿上の文書画像を読み取り画像データを出力する読取り工程と、前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程と、前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程と、前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程と、各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程とを有するものである。

【0011】第6の発明は、原稿上の文書画像を読み取り画像データを出力する読取り工程と、前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程と、前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程と、前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程と、

各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程とを有するコンピュータが読み出し可能なプログラムを記憶媒体に格納したものである。

【0012】第7の発明は、前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第1の指定工程と、前記第1の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定工程と、前記第1の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するものである。

【0013】第8の発明は、前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第1の指定工程と、前記第1の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定工程と、前記第1の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するコンピュータが読み出し可能なプログラムを記憶媒体に格納したものである。

【0014】第9の発明は、前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第1の表示工程と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示工程と、前記指示工程による修正指示に基づいて前記抽出工程により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程とを有するものである。

【0015】第10の発明は、前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第1の表示工程と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示工程と、前記指示工程による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程とを有するコンピュータが読み出し可能なプログラムを記憶媒体に格納したものである。

【0016】第11の発明は、前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示工程と、前記表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第2の指定工程と、前記第2の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定工程と、前記第2の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するものである。

【0017】第12の発明は、前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示工程と、前記

10

20

30

40

50

表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第 2 の指定工程と、前記第 2 の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第 2 の判定工程と、前記第 2 の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するコンピュータが読み出し可能なプログラムを記憶媒体に格納したものである。

【0018】

【発明の実施の形態】

〔第 1 実施形態〕図 1 は、本発明に係る第 1 実施形態を示すデータ処理装置の構成を説明するブロック図である。

【0019】図において、101 は CPU で、ROM 102 に格納されているプログラムにしたがってデータ処理装置全体の制御等を行う。102 は ROM で、CPU 101 が実行する後述する図 3、図 4、図 8、図 10 に示すフローチャートに基づくプログラムと本装置の制御プログラムと図形、イメージ、文字、表等が混在した画像を処理（文字認識処理、領域抽出処理等を含む）するためのプログラム等を格納する。

【0020】103 は RAM で、CPU 101 の主メモリ、ワークエリアなどとして機能し、また文書および画像等を記憶する。104 は外部記憶装置で、磁気ディスク（ハードディスク、フロッピーディスク、MO ディスク）等の記憶媒体から構成されている。105 はディスプレイで、データ処理装置の処理内容などを表示する。106 はキーボード、107 はポインティングデバイスで、それぞれを操作することによりデータ処理装置への入力及びアイコン指示を行う。

【0021】108 はイメージスキャナ（スキャナ）で、原稿等の画像を光学的に読み取り、画像データを RAM 103 または外部記憶装置 104 に出力する。109 は音声合成部で、文字コードが入力されると音声情報をスピーカ 110 に出力する。110 はスピーカで、音声合成部 109 等から得られた音声情報を音声として出力する。

【0022】また、このように構成されたデータ処理装置において、CPU 101 は ROM 102 に記憶された文字認識プログラムに基づき、スキャナ 108 等から画像として読み込まれた原稿を解析して、原稿に記載されている文章を文字コードとして認識することが可能である。

【0023】以下、図 2 を参照して、スキャナ 108 から原稿を読み込み、領域を抽出した時のディスプレイ 105 上に表示される原稿の表示状態と、読み込んだ原稿から領域を抽出する CPU 101 による処理について説明する。

【0024】図 2 は、図 1 に示したディスプレイ 105 上に表示される第 1 の原稿表示ウインドウを示す図であ

る。

【0025】図において、W1 は原稿表示ウインドウで、文字の大きさの異なる種々の文章とイメージとが混在した原稿 P1 が表示されている。1～8 は原稿 P1 から抽出された領域であり、各領域 1～8 は枠線 L1～L8 により囲まれ、他の領域と区別ができるように表示されている。

【0026】図 2 に示した原稿 P1 における各領域 1～8 の抽出処理の一例について説明する。

10 【0027】まず、入力画像である原稿 P1 を M×N 毎の小領域に分割し、その分割された小領域内の少なくとも 1 つの画素に画像が存在すれば、その小領域に画像があるものと判断して間引き画像を作成する。

【0028】次に、間引いた画像が連結する部分をまとめ小矩形を作っていく。また、縦横比が大きく短辺同士が接近している矩形は文字行である可能性が高いので、このような矩形を結合する。さらに、短辺の長さがほぼ同じで、ほぼ等間隔に並んでいる矩形の集合は、本文部である可能性が高いので、これら矩形の集合を 1 つの領域に結合する。その結果、領域 1～領域 8 が抽出される。

【0029】次に、抽出された領域に属性を付加する処理について説明する。領域 2 は領域の抽出過程（複数の小矩形を併合して作成された）から本文部である可能性が高いと判断され、“本文”の属性を付加する。同様に、領域 4、領域 6 にも“本文”の属性を付加する。

【0030】次に、残りの領域について、まず、文字がふくまれているかどうかを判断する。文字が含まれている領域は画像の有無を示す小矩形が周期的に現れるので、その性質を利用して文字が含まれている領域を選択できる。その結果、領域 1、領域 5、領域 7 は文字が含まれる領域として選択される。

【0031】さらに、これらが本文部の文字行の行幅より大きく、本文部の段とは異なる位置にあるという判定条件を満たせば“見出し”、本文部の文字行の行幅より大きい、本文部の後段と領域の幅がほぼ等しいという判定条件を満たせば“小見出し”、本文部の文字行の行幅より大きくなり、原稿の下端部や上端部に存在するという判定条件を満たせば“ページ”の属性を自動的に付加する。

【0032】その結果、領域 1 には“見出し”、領域 5 には“小見出し”、領域 7 には“ページ”の属性が自動的に付加される。残りの領域について領域の大きさが小さければ“ノイズ”とし、それ以外は絵や写真であるとして“イメージ”とする。領域 3 は、領域が十分大きいので“イメージ”の属性を付加し、領域 8 には“ノイズ”の属性を付加する。

【0033】このように、抽出された領域に上記のような属性を自動的に付加することにより、各領域が原稿に対してどのような属性を持っているのかを容易に判別す

ることができるようになる。

【0034】以下、本実施形態の特徴的構成について図1等を参照して説明する。

【0035】上記のように構成された、原稿上の文書画像を読み取り画像データを出力する読取り手段（スキャナ108）と、前記読取り手段により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出手段（領域抽出処理を行うCPU101）と、前記抽出手段により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加手段（CPU101）と、前記付加手段により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識手段（文字認識処理を行うCPU101）と、各領域に対して前記認識手段により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源（RAM103、外部記憶装置104）に保存する保存手段（RAM103、外部記憶装置104を制御するCPU101）とを有するので、登録すべき原稿画像中から音声データをキーとして検索可能な文字情報領域を切り出して登録情報と共に自動保存でき、ユーザによる登録画像の検索情報の入力処理を軽減し、当該登録画像を連想するにふさわしい有効な検索情報を簡便に入力できる。

【0036】また、文字情報を音声として出力する出力手段（音声合成部109、スピーカ110）と、前記メモリ資源に保存された画像データに対して検索すべき領域の属性を指定する第1の指定手段（キーボード106、ポインティングデバイス107）と、前記第1の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定手段（後述する図4に示すフローチャートに基づく処理を行うCPU101）と、前記第1の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第1の制御手段（音声合成部109を制御するCPU101）とを有するので、属性が付加されている領域をオペレータが指定するという簡便な操作で、画像データ全体のイメージを表示させるための多大な処理時間を必要とすることなく、登録した画像内容を音声で効率よく確認させることができる。

【0037】図3は、本発明に係るデータ処理装置における第1のデータ処理手順の一例を示すフローチャートであり、図1に示したスキャナ108から原稿を読み取り、外部記憶装置104に保存する手順に対応する。なお、(1)～(5)は各ステップを示す。

【0038】まず、イメージスキャナ108等から光電変換することにより例えば原稿P1の画像を入力し

(1)、該入力された画像データを上述した領域抽出手段に従ってそれぞれの領域を抽出し(2)、分割された各領域1～8毎に属性（本文、見出し、小見出し、ページ等）を付加する(3)。

【0039】次に、文字が含まれる属性の領域について公知の文字認識処理を実行し、該認識結果により得られる文字情報を該領域に対して付加し(4)、得られた領域の位置や大きさ等の情報、領域の属性、領域の文字情報が存在すれば文字情報等を外部記憶装置104等に保存し(5)、処理を終了する。

【0040】この処理を繰り返すことで、スキャナ108から読み込まれた複数の原稿が外部記憶装置104に領域情報の組み合わせとして記憶される。以後、指定入力される属性を検索キーとして記憶された画像情報を要部検索することが可能となる。

【0041】図4は、本発明に係るデータ処理装置における第2のデータ処理手順の一例を示すフローチャートであり、図1に示した外部記憶装置104に記憶された原稿情報の検索手順に対応する。なお、(1)～(3)は各ステップを示す。

【0042】まず、キーボード106、ポインティングデバイス107によりどの属性に基づいて検索を行うかを指定し、該指定された属性と同一の属性を持つ領域を外部記憶装置104に保存された原稿情報の中から検索して該当する領域を選択し(1)、該選択された領域が存在するかどうかを判断し(2)、該当する領域が存在しないと判断された場合は処理を終了する。

【0043】一方、ステップ(2)で該当する領域が存在すると判断された場合は、選択された領域の文字情報の文字コードを外部記憶装置104から読み出し音声合成部109に送って音声合成を行い、スピーカ110から音声として読み上げ出力し(3)、処理を終了する。

【0044】これにより、例えばステップ(1)において属性を“見出し”と指定して、図2に示した原稿P1の検出を行うと領域1が選択される。その結果、図2に示した領域1に付加された文字情報である「音声合成技術について」がスピーカ110より読み上げられることになる。

【0045】以下、本実施形態の特徴的構成について図3、図4等を参照して説明する。

【0046】上記のように構成された、原稿上の文書画像を読み取り画像データを出力する読取り工程（図3のステップ(1)）と、前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程（図3のステップ(2)）と、前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程（図3のステップ(3)）と、前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程（図3のステップ(4)）と、各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程（図3のステップ(4)、(5)）とを有するので、登録すべき

原稿画像中から音声キーとして検索可能な文字情報領域を切り出して登録情報と共に自動保存でき、ユーザによる登録画像の検索情報の入力処理を軽減し、当該登録画像を連想するにふさわしい有効な検索情報を簡便に入力できる。

【0047】また、前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第1の指定工程（不図示のステップ）と、前記第1の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定工程（図4のステップ（2））と、前記第1の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程（図4のステップ（3））とを有するので、属性が付加されている領域をオペレータが指定するという簡便な操作で、画像データ全体のイメージを表示させるための多大な処理時間を必要とすることなく、登録した画像内容を音声で効率よく確認させることができる。

【0048】以上説明したように、本実施形態により、属性を付加して原稿を保存した後、適時オペレータが所望の特定の属性を指定するという簡単な操作で、その属性で記憶されている領域内の原稿情報を読み上げることができ、原稿の内容をイメージとしてディスプレイ105上に再現することなく簡便に確認することができる効果がある。

【0049】なお、画像の読み込みはイメージスキャナ（スキャナ108）に限るものではなく、外部記憶装置104等にあらかじめ記憶されたものであってもよい。また、領域情報はRAM103に記憶するよう構成してもよい。また、通信手段を用いて遠隔地にある他のデータ処理装置や記憶装置等に記憶されている原稿を読み込めるよう構成してもよい。

【0050】更に、音声合成手段については音声合成部109に限るものではなく、供給されるプログラムを読み込みCPU101が実行することにより音声情報を合成するように構成されていても問題ない。

【0051】また、領域抽出方法や属性付加方法については同様の効果が得られるものであれば本実施形態で説明した方法に限るものではなく、属性の種類についても本実施形態で説明したものに限らないことは言うまでもない。

【0052】〔第2実施形態〕上記第1実施形態では、データ処理装置が原稿画像を解析して各領域を抽出して適正な属性を自動的に付加し、文字を含む領域には文字認識された文字情報を領域に付加して外部記憶装置104に記憶する場合を説明したが、データ処理装置で自動的に求められた領域、属性、文字情報をユーザが修正できるように構成しても良い。以下、その実施形態について説明する。

【0053】なお、本実施形態におけるデータ処理装置は図1に示した第1実施形態と同一の構成を有してい

る。

【0054】以下、図5を参照して、データ処理装置により抽出された領域をユーザが修正する場合にディスプレイ105上に表示される第2の原稿表示ウインドウと該画面におけるグラフィカルユーザインタフェースについて説明する。

【0055】図5は、図1に示したディスプレイ105上に表示される第2の原稿表示ウインドウの一例を説明する図である。なお、図2と同一のものには同一の符号を付してある。また、該原稿表示ウインドウに表示される原稿は図2に示した原稿と同一であるが、領域の抽出結果が異なる場合を示している。

【0056】図において、W2は原稿表示ウインドウで、文字の大きさの異なる種々の文章とイメージとが混在した原稿P1が表示されている。1a、1b、2～8は原稿P1から抽出された領域であり、各領域1a、1b、2～8はそれぞれ枠線L1a、L1b、L2～L8により囲まれ、他の領域と区別できるように表示されている。なお、領域1a、領域1bは見出しである文章が「音声合成技術について」と「て」との二つの部分に分割されて抽出されている場合に対応する。なお、図示していないがこのウインドウW2上では抽出された領域に対応する属性および文字情報を表示し、これらの情報を確認できるように表示されているものとする。

【0057】この原稿表示ウインドウW2上において、例えば領域1bを選択する場合は、ポインティングデバイス107等を使用して領域1b内を指示してクリックすれば達成され、この時選択された領域1bは枠線L1bを他の領域より太く表示される。なお、選択されているのが分かりやすくなる表示がより良いのは言うまでもない。

【0058】また、例えば領域1aの大きさを変更する場合には、ポインティングデバイス107で領域1aの境界（枠線L1a）付近を指示し、ドラッグ操作により枠線L1aをつかんで所望の位置まで移動してドラッグすることにより達成される。また、必要となる領域が自動抽出できなかったために、新たな領域を作成しなければならない場合には、ポインティングデバイス107を操作することにより領域を新たに作成することができる。

【0059】例えば、図に示した領域1aと領域1bに別れて抽出されてしまった見出し部を1つの領域にまとめる場合には、まず、領域1bを選択して削除を行い、領域1bに含まれていた「て」が領域1aの枠線L1a内に含まれるようにポインティングデバイス107によりドラッグ操作して領域1aの枠線L1aの大きさを変更させることにより、変更された領域1a、すなわち領域1に見出し全体が含まれるように枠線L1が見出し全体を囲むように表示して、領域1の領域情報、属性等を変更することができる。



【0060】図5に示したように、ユーザのポインティングデバイス107による簡単な操作指示により原稿表示ウインドウW2上で、領域の追加、削除、変更等の修正ができる。

【0061】以下、図6を参照して、領域に付加された属性を編集する場合にディスプレイ105上に表示されるウインドウを説明する。

【0062】図6は、図1に示したディスプレイ105上に表示される属性修正ウインドウの一例を説明する図である。

【0063】図において、W3は属性修正ウインドウで、属性を修正するために、プルダウンメニュー等に登録されている属性修正項目をポインティングデバイス107等のダブルクリック操作により開かれる画面であり、OKボタンB1とキャンセルボタンB2と属性修正ボックスBOX1とから構成されている。

【0064】また、属性修正ボックスBOX1は属性を「本文」と修正する場合に選択する項目Z1、「見出し」と修正する場合に選択する項目Z2、「小見出し」と修正する場合に選択する項目Z3、「ページ」と修正する場合に選択する項目Z4、「イメージ」と修正する場合に選択する項目Z5、「ノイズ」と修正する場合に選択する項目Z6から構成されている。

【0065】なお、属性修正ウインドウW3上では、ポインティングデバイス107を操作して属性修正ボックスBOX1内の項目Z1～Z6を選択して、OKボタンB1をクリックすることにより属性を修正することができるように構成されている。

【0066】例えば、図5に示した領域1aの属性が“本文”になっている場合は、領域1aを選択し、属性修正ウインドウW3上で正しい属性を選択することにより修正する。なお、属性修正ウインドウW3を属性名が表示された領域の近傍に表示したり、属性毎の領域の色を付けて表示して、領域の属性が判別し易いようにしてもよい。

【0067】このように、属性修正ウインドウW3上でポインティングデバイス107等を用いて操作することにより、自動的に付加された領域の属性が実際の属性と異なっていた場合に、ユーザの選択で修正することができる。

【0068】以下、図7を参照して、領域に付加された文字情報を編集する場合にディスプレイ105上に表示されるウインドウを説明する。

【0069】図7は、図1に示したディスプレイ105上に表示される文字情報ウインドウの一例を説明する図である。

【0070】図において、W4は文字情報ウインドウで、文字情報を修正するために、プルダウンメニュー等に登録されている文字情報修正項目をポインティングデバイス107等のダブルクリック操作により開かれる画

面であり、終了ボタンB3と文字情報修正ボックスBOX2からなり、文字情報修正ボックスBOX2内に文字情報MJが表示されている。なお、文字情報MJ中の「会」の部分が選択されている。また、終了ボタンB3をクリックすることにより文字情報の修正を終了することができる。

【0071】なお、この図では図5に示した領域1の文字情報をコンピュータによる文字認識機能を使用して作成したもので、「音声合成技術について」と認識するところを「音声合成技術について」と誤認識してしまった状態を表示している場合に対応する。

【0072】このように誤認識が発生した場合に、文字情報ウインドウW4上において、キーボード106またはポインティングデバイス107を操作して、文字情報修正ボックスBOX2内の文字情報を書き換えることにより、文字情報を修正することができるように構成されている。

【0073】例えば、図5に示したように領域1が「・・・会成・・・」となっている場合に、領域1を選択し、文字情報ウインドウW4を表示して、「会」を「合」に修正する。

【0074】このように、文字情報ウインドウW4上でポインティングデバイス107等を用いて操作することにより、自動的に認識された領域の文字情報が実際の文字情報と異なっていた場合に、ユーザが直接入力することで自在に修正又は変更することができる。

【0075】以下、本実施形態の特徴的構成について図1、図5、図6、図7等を参照して説明する。

【0076】上記のように構成された、前記抽出手段により抽出された矩形領域および前記認識手段により認識された文字情報を表示部（ディスプレイ105）に表示させる第1の表示制御手段と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示手段（原稿表示ウインドウW2、属性修正ウインドウW3、文字情報ウインドウW4上でキーボード106、ポインティングデバイス107で指示）と、前記指示手段による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識手段により認識された文字情報を修正する編集手段（後述する図8に示すフローチャートに基づき処理するCPU101）とを有するので、抽出された矩形領域またはその属性データ、文字情報に不具合があった場合でも、矩形領域、属性データ、文字情報の修正、削除等をオペレータが簡単な操作だけで行うことができる。

【0077】以下、領域情報の保存に関する他の実施形態を図8のフローチャートを用いて説明する。

【0078】図8は、本発明に係るデータ処理装置における第3のデータ処理手順の一例を示すフローチャートである。なお、（1）～（8）は各ステップを示す。

【0079】まず、第1実施形態と同様にスキャナ10

8で原稿の画像を読み込み(1)、読み込まれた原稿の領域を上記同様に抽出し(2)、該抽出された領域に属性を自動的に付加し(3)、文字情報を付加する

(4)。なお、この時点で、ディスプレイ105に抽出された領域と属性、文字情報を表示し、これらの情報を確認可能とする。

【0080】次に、抽出された領域の編集を図5に示した原稿表示ウインドウW2上で行う(5)。次に、各領域の属性の修正を図6に示した属性修正ウインドウW3上で行う(6)。

【0081】次に、図7に示した文字情報ウインドウW4上で文字情報の編集を行い(7)、原稿をステップ(5)～(7)において編集、修正された領域、属性、文字情報の組み合わせで外部記憶装置104に保存して(8)、処理を終了する。

【0082】以下、本実施形態の特徴的構成について図8等を参照して説明する。

【0083】上記のように構成された、前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第1の表示工程(不図示のステップ)と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示工程(不図示のステップ)と、前記指示工程による修正指示に基づいて前記抽出工程により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程(ステップ(5)～(7))とを有するので、抽出された矩形領域またはその属性データ、文字情報に不具合があった場合でも、矩形領域、属性データ、文字情報の修正、削除等をオペレータが簡単な操作だけで行うことができる。

【0084】以上説明したように、本実施形態により領域情報や属性、文字情報を編集、修正可能とすることで、領域抽出不良、文字認識不良に伴う領域、該領域の属性、領域内の文字をユーザが意図する内容に修正、変更、削除することができる。

【0085】なお、本実施形態では図8のステップ(1)～(3)で、あらかじめ領域情報を作成した後に編集、修正作業を行っているが、全て、あるいは一部の領域情報を手動で作成するように構成されていてもよく、また、図8のステップ(5)～(7)の順序は、本実施形態に示した順序に限ったことではない。

【0086】〔第3実施形態〕上記第2実施形態では、自動的に決定された領域、属性、文字情報を必要がある場合修正してから外部記憶装置104に保存する場合を説明したが、外部記憶装置104に複数保存されている各原稿の領域が、指定された属性と同一の属性を持つかどうかを検索して、該当する領域の複数の文字情報を順次読み上げるよう構成しても良い。以下、その実施形態について説明する。

【0087】なお、本実施形態におけるデータ処理装置

は、図1に示した第1実施形態と同一の構成を有している。

【0088】以下、図9を参照して、外部記憶装置104に複数記憶されている原稿を検索する場合にディスプレイ105上に表示される検索ウインドウと該画面におけるグラフィカルユーザインタフェースについて説明する。

【0089】図9は、図1に示したディスプレイ105上に表示される検索設定ウインドウの一例を説明する図である。なお、図6と同一のものには同一の符号を付してある。

【0090】図において、W5は検索設定ウインドウで、キーボード106またはポインティングデバイス107により検索の実行指示がなされた場合に開かれるウインドウで、属性指定部BOX3と原稿指定部BOX4とから構成されている。属性指定部BOX3は、属性設定項目Z1～Z6が表示されている。原稿指定部BOX4は、外部記憶装置104などに記憶されている原稿1～5をアイコンF1～F5を用いて表示しており、反転表示されるアイコンF2、F3、F5が検索対象として指定された原稿に対応する。

【0091】上記のように構成された検索設定ウインドウW5上において、キーボード106またはポインティングデバイス107を操作して、属性指定部BOX3内の属性設定項目Z1～Z6を選択することにより検索する属性を設定することができ、また、原稿指定部BOX4内のアイコンF1～F5を選択することにより検索する原稿を選択することができるように構成されている。なお、B2はキャンセルボタンで、このキャンセルボタンB2を押すことにより検索処理を中断させることができる。また、B1はOKボタンで、該OKボタンB1が押下されると検索を開始する。

【0092】なお、原稿指定部BOX4は、本実施形態に限定されるものではなく、名称の一覧を表示したり、先頭ページの縮小画像を表示したり、原稿内の文字情報の一部を原稿の識別に使うように構成されていてもよい。

【0093】このように、検索設定ウインドウW5上で、ポインティングデバイス107等を操作することにより、検索する属性および原稿を簡単に設定することができる。

【0094】以下、本実施形態の特徴的構成について図1、図9等を参照して説明する。

【0095】上記のように構成された、文字情報を音声として出力する出力手段(音声合成部109、スピーカ110)と、前記メモリ資源に保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示制御手段(検索ウインドウW5を表示するCPU101)と、前記表示部に表示される原稿画像および属性項目から所望の原稿画像および属性項

目を指定する第2の指定手段（キーボード106、ポインティングデバイス107）と、前記第2の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定手段（後述する図10に示すフローチャートに基づき判定するCPU101）と、前記第2の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第2の制御手段（音声合成部109を制御するCPU101）とを有するので、登録した画像を一覧表示でき、少ない操作で登録した画像の内容を次々と音声で効率よく検索処理

【0096】以下、図10を参照して第4のデータ処理手順について説明する。

【0097】図10は、本発明に係るデータ処理装置における第4のデータ処理手順の一例を示すフローチャートである。なお、(1)～(7)は各ステップを示す。

【0098】まず、検索設定ウィンドウ3の属性指定部BOX3上で検索する属性を指定し(1)、原稿指定部BOX4上で検索する原稿を指定し(2)、該指定された原稿、例えばアイコンF2、F3、F5が全て検索されたかどうかを判断し(3)、全ての原稿が検索されたと判断された場合は、処理を終了する。

【0099】一方、全ての原稿が検索されていないと判断された場合は、ステップ(4)以降に進む。なお、ステップ(4)～(6)は第1実施形態と同様で、指定した属性の領域が存在すればその文字情報をスピーカ110より読み上げ出力させる。

【0100】次に、オペレータからの処理を中断する指示が検知されたかどうか、すなわち図9に示したキャンセルボタンB2が押されたかどうかを判断し(7)、押されていないと判断した場合は、ステップ(3)に戻り、一方、押されたと判断した場合は、処理を終了する。

【0101】以上説明したように、本実施形態により検索すべき属性や文書をユーザが細かく設定可能としたため、該検索処理を中断することにより、必要以上の原稿情報の無駄な読み上げを防止できる効果がある。

【0102】なお、本実施形態では、検索条件として属性と原稿だけを指定できるように構成されているが、検索条件に文字情報を付加し、指定属性の領域であり、かつ、文字情報に指定文字列を含む領域、あるいは指定属性の領域で、または指定文字列を含む文字情報をもつ領域を検索できるように検索条件を指定するように構成されていてもよい。

【0103】また、検索されている原稿名も読み上げるように構成されていたり、選択された領域が存在しない原稿について、選択領域が存在しない旨を音声で出力するように構成されていてもよい。

【0104】さらに、本実施形態では中断の検知タイミングは、最初に検索する原稿の処理が終了した時点で

われているが、このタイミングに限るものではない。

【0105】以下、本実施形態の特徴的構成について図10等を参照して説明する。

【0106】上記のように構成された、前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示工程（不図示のステップ）と、前記表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第2の指定工程（図10のステップ(1)）、

(2)と、前記第2の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定工程（図10のステップ(5)）と、前記第2の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程（図10のステップ(6)）とを有するので、登録した画像を一覧表示でき、少ない操作で登録した画像の内容を次々と音声で効率よく検索処理できる。

【0107】以下、図11に示すメモリマップを参照して本発明に係るデータ処理装置で読み出し可能なデータ処理プログラムの構成について説明する。

【0108】図11は、本発明に係るデータ処理装置で読み出し可能な各種データ処理プログラムを格納する記憶媒体のメモリマップを説明する図である。

【0109】なお、特に図示しないが、記憶媒体に記憶されるプログラム群を管理する情報、例えばバージョン情報、作成者等も記憶され、かつ、プログラム読み出し側のOS等に依存する情報、例えばプログラムを識別表示するアイコン等も記憶される場合もある。

【0110】さらに、各種プログラムに従属するデータも上記ディレクトリに管理されている。また、各種プログラムをコンピュータにインストールするためのプログラムや、インストールするプログラムが圧縮されている場合に、解凍するプログラム等も記憶される場合もある。

【0111】本実施形態における図3、図4、図8、図10に示す機能が外部からインストールされるプログラムによって、ホストコンピュータにより遂行されていてもよい。そして、その場合、CD-ROMやフラッシュメモリやFD等の記憶媒体により、あるいはネットワークを介して外部の記憶媒体から、プログラムを含む情報群を出力装置に供給される場合でも本発明は適用されるものである。

【0112】以上のように、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、本発明の目的が達成されることは言うまでもない。

【0113】この場合、記憶媒体から読み出されたプロ

グラムコード自体が本発明の新規な機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0114】プログラムコードを供給するための記憶媒体としては、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROM、EEPROM等を用いることができる。

【0115】また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0116】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPU等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0117】

【発明の効果】以上説明したように、本発明に係る第1の発明によれば、原稿上の文書画像を読み取り画像データを出力する読取り手段と、前記読取り手段により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出手段と、前記抽出手段により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加手段と、前記付加手段により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識手段と、各領域に対して前記認識手段により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存手段とを有するので、登録すべき原稿画像中から音声データをキーとして検索可能な文字情報領域を切り出して登録情報と共に自動保存でき、ユーザによる登録画像の検索情報の入力処理を軽減し、当該登録画像を連想するにふさわしい有効な検索情報を簡便に入力できる。

【0118】第2の発明によれば、文字情報を音声として出力する出力手段と、前記メモリ資源に記憶された画像データに対して検索すべき領域の属性を指定する第1の指定手段と、前記第1の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定手段と、前記第1の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第1の制御手段とを有するので、属性が付加されている領域をオペレータが指定するという簡便な操作

で、画像データ全体のイメージを表示させるための多大な処理時間を必要とすることなく、登録した画像内容を音声で効率よく確認させることができる。

【0119】第3の発明によれば、前記抽出手段により抽出された矩形領域および前記認識手段により認識された文字情報を表示部に表示させる第1の表示制御手段と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示手段と、前記指示手段による修正指示に基づいて前記抽出手段により抽出された矩形領域または前記認識手段により認識された文字情報を修正する編集手段とを有するので、抽出された矩形領域またはその属性データ、文字情報に不具合があった場合でも、矩形領域、属性データ、文字情報の修正、削除等をオペレータが簡単な操作だけで行うことができる。

【0120】第4の発明によれば、文字情報を音声として出力する出力手段と、前記メモリ資源に保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示制御手段と、前記表示部に表示される原稿画像および属性項目から所望の原稿画像および属性項目を指定する第2の指定手段と、前記第2の指定手段により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定手段と、前記第2の判定手段の判定結果に基づいて前記出力手段による音声読み上げを制御する第2の制御手段とを有するので、登録した画像を一覧表示でき、少ない操作で登録した画像の内容を次々と音声で効率よく検索処理できる。

【0121】第5、第6の発明によれば、原稿上の文書画像を読み取り画像データを出力する読取り工程と、前記読取り工程により読み取られた前記画像データを解析して前記原稿画像中から任意サイズの矩形領域を抽出する抽出工程と、前記抽出工程により抽出される各矩形領域の形態に基づいて異なる属性データを付加する付加工程と、前記付加工程により属性データが付加された各矩形領域内の画像データを解析して文字情報を認識する認識工程と、各領域に対して前記認識工程により認識された文字情報および前記属性データを付加して前記画像データとともにメモリ資源に保存する保存工程とを有するので、登録すべき原稿画像中から音声データをキーとして検索可能な文字情報領域を切り出して登録情報と共に自動保存でき、ユーザによる登録画像の検索情報の入力処理を軽減し、当該登録画像を連想するにふさわしい有効な検索情報を簡便に入力できる。

【0122】第7、第8の発明によれば、前記保存工程により保存された画像データに対して検索すべき領域の属性を指定する第1の指定工程と、前記第1の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第1の判定工程と、前記第1の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するので、オペレータが

属性が付加されている領域を指定するという簡便な操作で、画像データ全体のイメージを表示させるための多大な処理時間を必要とすることなく、登録した画像内容を音声で効率よく確認させることができる。

【0123】第9、第10の発明は、前記抽出工程により抽出された矩形領域および前記認識工程により認識された文字情報を表示部に表示させる第1の表示工程と、前記表示部により表示された矩形領域または文字情報に対する修正指示を行う指示工程と、前記指示工程による修正指示に基づいて前記抽出工程により抽出された矩形領域または前記認識工程により認識された文字情報を修正する編集工程とを有するので、抽出された矩形領域またはその属性データ、文字情報に不具合があった場合でも、矩形領域、属性データ、文字情報の修正、削除等をオペレータが簡単な操作だけで行うことができる。

【0124】第11、第12の発明は、前記保存工程により保存された原稿画像の一覧および検索対象とする属性項目の一覧を前記表示部に表示させる第2の表示工程と、前記表示部に表示される原稿画像及び属性項目から所望の原稿画像及び属性項目を指定する第2の指定工程と、前記第2の指定工程により指定された属性に基づく領域が前記メモリ資源から選択されたかどうかを判定する第2の判定工程と、前記第2の判定工程の判定結果に基づいて文字情報を音声として出力する出力工程とを有するので、登録した画像を一覧表示でき、少ない操作で登録した画像の内容を次々と音声で効率よく検索処理できる。

【0125】従って、原稿画像の領域に付加されている文字情報を読み上げて、原稿画像全体のイメージを表示させるための多大な処理時間を必要とすることなく、オペレータの特定の属性および原稿画像を指定するという簡単な操作だけで、短時間の内に原稿画像の内容を音声でオペレータに確認させることができる。

【図面の簡単な説明】

\* 【図1】本発明に係る第1実施形態を示すデータ処理装置の構成を説明するブロック図である。

【図2】図1に示したディスプレイ上に表示される第1の原稿表示ウィンドウの一例を示す図である。

【図3】本発明に係るデータ処理装置における第1のデータ処理手順の一例を示すフローチャートである。

【図4】本発明に係るデータ処理装置における第2のデータ処理手順の一例を示すフローチャートである。

【図5】図1に示したディスプレイ上に表示される第2の原稿表示ウィンドウの一例を説明する図である。

【図6】図1に示したディスプレイ上に表示される属性修正ウィンドウの一例を説明する図である。

【図7】図1に示したディスプレイ上に表示される文字情報ウィンドウの一例を説明する図である。

【図8】本発明に係る原稿処理装置における第3のデータ処理手順の一例を示すフローチャートである。

【図9】図1に示したディスプレイ上に表示される検索設定ウィンドウの一例を説明する図である。

【図10】本発明に係るデータ処理装置における第4のデータ処理手順の一例を示すフローチャートである。

【図11】本発明に係るデータ処理装置で読み出し可能な各種データ処理プログラムを格納する記憶媒体のメモリマップを説明する図である。

【符号の説明】

101 CPU

102 ROM

103 RAM

104 外部記憶装置

105 ディスプレイ

106 キーボード

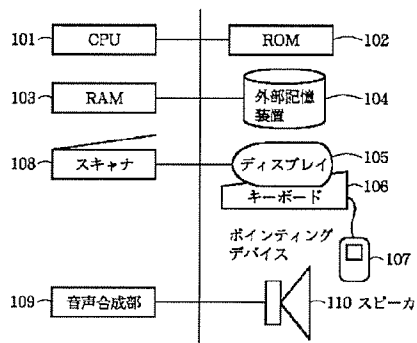
107 ポインティングデバイス

108 スキャナ

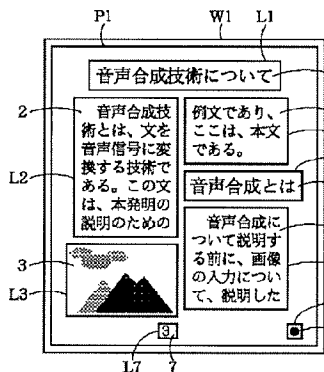
109 音声合成部

110 スピーカ

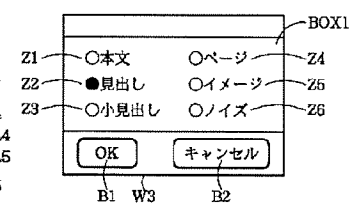
【図1】



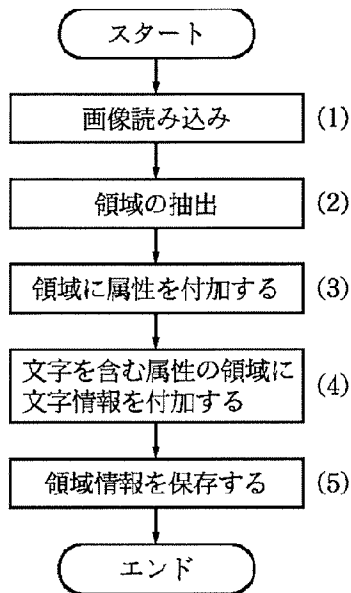
【図2】



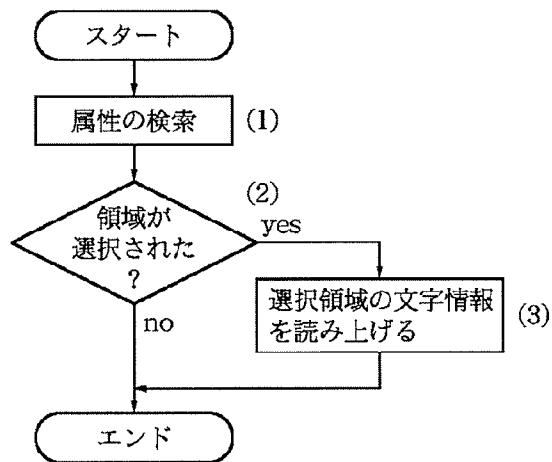
【図6】



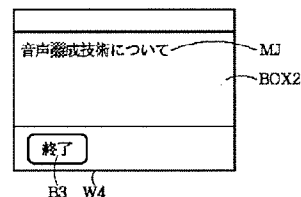
【図 3】



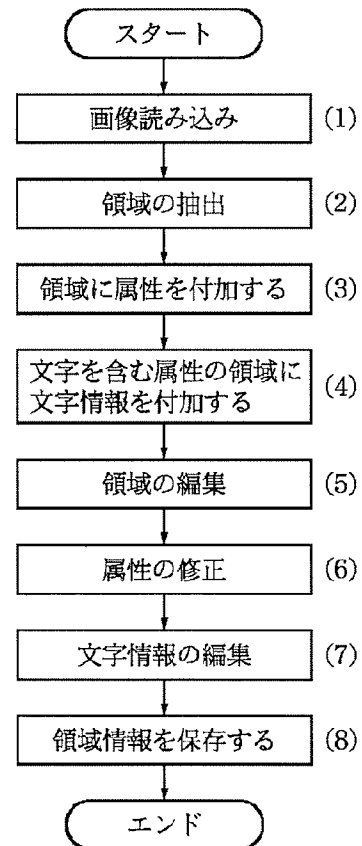
【図 4】



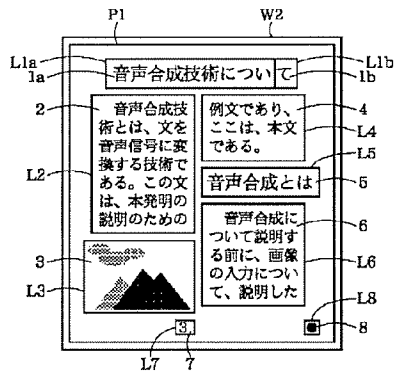
【図 7】



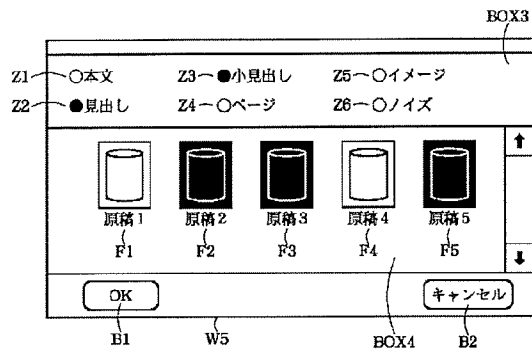
【図 8】



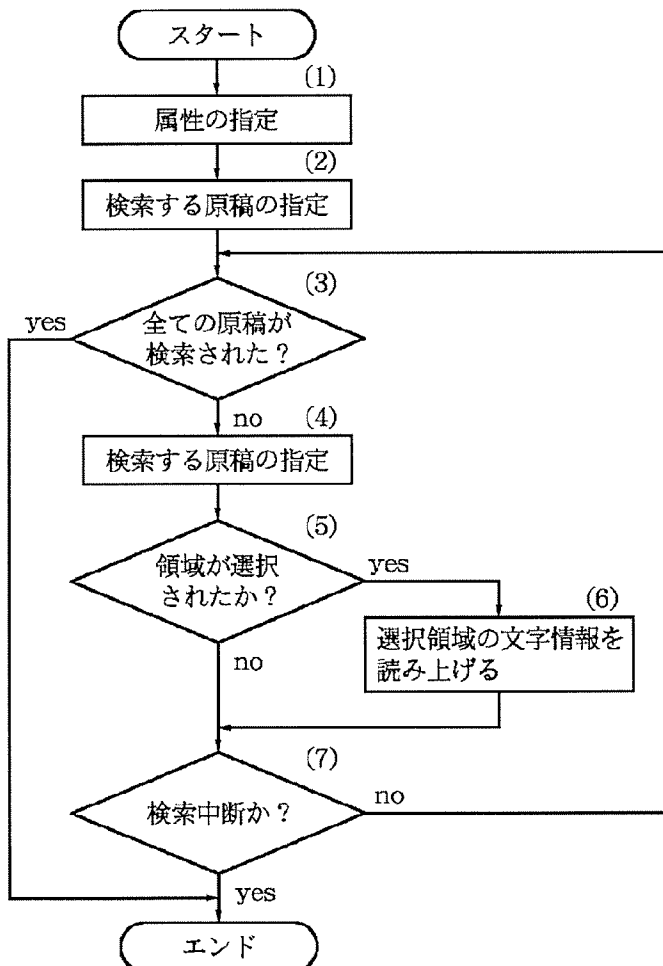
【図 5】



【図 9】



【図 10】



【図 11】

| FD/CD-ROM等の記憶媒体 |                                  |
|-----------------|----------------------------------|
| ディレクトリ情報        |                                  |
| 第1のデータ処理プログラム   | 図3に示すフローチャートのステップに対応するプログラムコード群  |
| 第2のデータ処理プログラム   | 図4に示すフローチャートのステップに対応するプログラムコード群  |
| 第3のデータ処理プログラム   | 図8に示すフローチャートのステップに対応するプログラムコード群  |
| 第4のデータ処理プログラム   | 図10に示すフローチャートのステップに対応するプログラムコード群 |

記憶媒体のメモリマップ

(11) Japanese Patent Laid-Open No. 11-85457

(43) Laid-Open Date: March 30, 1999

(21) Application No. 9-246351

(22) Application Date: September 11, 1997

(71) Applicant: CANON KABUSHIKI KAISHA

(72) Inventor: Hiroaki Ikeda

(54) [Title of the Invention] DATA PROCESSING  
APPARATUS, DATA PROCESSING METHOD FOR DATA PROCESSING  
APPARATUS, AND STORAGE MEDIUM STORING COMPUTER-READABLE  
PROGRAM

(57) [Abstract]

[Problems to be Solved]

To allow an operator to confirm the contents of an  
original image in voice in a short time without  
displaying the entire image of the original image.

[Solution]

A CPU 101 analyzes a read original image and  
extracts the image in a rectangular area of an optional  
size, adds different attribute data based on the form  
of each of the extracted rectangular areas, analyzes  
the image data in the rectangular area to which the  
attribute data is added and recognizes character  
information, adds the character information and the  
attribute data and stores the result with the image



data in an external storage device 104, specifies the original image and attribute to be retrieved for the stored original image, determines whether or not an area based on the specified original image and attribute has been selected from the external storage device 104, and based on the determination result, a speech synthesis unit 109 controls voice-reading through a speaker 110 of the character information added to the area information.

[Claims for the Patent]

[Claim 1]

A data processing apparatus, comprising:

read means for reading a document image on an original and outputting image data;

extraction means for analyzing the image data read by said read means, and extracting a rectangular area of an optional size from the original image;

addition means for adding attribute data different based on a form of each rectangular area extracted by said extraction means;

recognition means for analyzing image data in each rectangular area to which attribute data is added by said addition means, and recognizing character information; and

storage means for adding the character information recognized by said recognition means and the attribute data to each area, and storing resultant data with the image data in memory resources.

[Claim 2]

The data processing apparatus according to claim 1, further comprising:

output means for voice-outputting character information;

first specification means for specifying an attribute of an area to be retrieved for image data stored in the memory resources;

first determination means for determining whether or not an area based on the attribute specified by said first specification means is selected from the memory resources; and

first control means for controlling voice-reading by said output means based on a determination result by said first determination means.

[Claim 3]

The data processing apparatus according to claim 1, further comprising:

first display control means for displaying the rectangular area extracted by said extraction means and the character information recognized by said recognition means on a display unit;

instruction means for issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and

edition means for correcting the rectangular area extracted by said extraction means or the character information recognized by said recognition means in response to the correct instruction from said instruction means.

[Claim 4]

The data processing apparatus according to claim 3, further comprising;

output means for voice-outputting character information;

second display control means for displaying on the display unit a list of original images stored in the memory resources and a list of attribute items to be retrieved;

second specification means for specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit;

second determination means for determining whether or not an area based on the attribute specified by said second specification means is selected from the memory resources; and

second control means of controlling voice-reading by said output means based on a determination result by said second determination means.

[Claim 5]

A data processing method for a data processing apparatus, comprising:

a reading step of reading a document image on an original and outputting image data;

an extracting step of analyzing the image data read in said reading step, and extracting a rectangular area of an optional size from the original image;

an adding step of adding attribute data different based on a form of each rectangular area extracted in said extracting step;

a recognizing step of analyzing image data in each rectangular area to which attribute data is added in said adding step, and recognizing character information; and

a storing step of adding the character information recognized in said recognizing step and the attribute data to each area, and storing resultant data with the image data in memory resources.

[Claim 6]

A storage medium storing a computer-readable program, comprising:

a reading step of reading a document image on an original and outputting image data;

an extracting step of analyzing the image data read in said reading step, and extracting a rectangular area of an optional size from the original image;

an adding step of adding attribute data different based on a form of each rectangular area extracted in said extracting step;

a recognizing step of analyzing image data in each rectangular area to which attribute data is added in said adding step, and recognizing character information; and

a storing step of adding the character information recognized in said recognizing step and the attribute data to each area, and storing resultant data with the image data in memory resources.

[Claim 7]

The data processing method for a data processing apparatus according to claim 5, comprising:

a first specifying step of specifying an attribute of an area to be retrieved for the image data stored in said storing step;

a first determining step of determining whether or not an area based on the attribute specified in said first specifying step is selected from the memory resources; and

an outputting step of voice-outputting character information based on the determination result in said first determining step.

[Claim 8]

The storage medium storing a computer-readable program according to claim 6, comprising:

a first specifying step of specifying an attribute of an area to be retrieved for the image data stored in said storing step;

a first determining step of determining whether or not an area based on the attribute specified in said first specifying step is selected from the memory resources; and

an outputting step of voice-outputting character information based on the determination result in said first determining step.

[Claim 9]

The data processing method for a data processing apparatus according to claim 5, further comprising:

- a first display step of displaying the rectangular area extracted in said extracting step and the character information recognized in said recognizing step on a display unit;

- an instructing step of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and

- an editing step of correcting the rectangular area extracted in said extracting step or the character information recognized in said recognizing step in response to the correct instruction in said instructing step.

[Claim 10]

The storage medium storing a computer-readable program according to claim 6, further comprising:

- a first display step of displaying the rectangular area extracted in said extracting step and the character information recognized in said recognizing step on a display unit;

- an instructing step of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and

- an editing step of correcting the rectangular area extracted in said extracting means or the character information recognized in said recognizing step in

response to the correct instruction in said instructing step.

[Claim 11]

The data processing method for a data processing apparatus according to claim 9, further comprising;

a second display step of displaying on the display unit a list of original images stored in said storing step and a list of attribute items to be retrieved;

a second specifying step of specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit;

a second determining step of determining whether or not an area based on the attribute specified in said second specifying step is selected from the memory resources; and

an output step of voice-outputting character information based on a determination result in said second determining step.

[Claim 12]

The storage medium storing a computer-readable program according to claim 10, further comprising;

a second display step of displaying on the display unit a list of original images stored in said storing step and a list of attribute items to be retrieved;

a second specifying step of specifying a desired original image and attribute item from the original



images and attribute items displayed on the display unit;

a second determining step of determining whether or not an area based on the attribute specified in said second specifying step is selected from the memory resources; and

an output step of voice-outputting character information based on a determination result in said second determining step.

[Detailed Description of the Invention]

[0001]

[Field of the Invention]

The present invention relates to a data processing apparatus capable of adding, a specific attribute for each area extracted after reading an original image, storing areas and voice-outputting character information in the stored area, a data processing method for the data processing apparatus, and a storage medium storing a computer-readable program.

[0002]

[Conventional Art]

In the conventional data processing apparatus having an electronic filing function, when an original image input from a scanner etc. is stored in a storage device, a user-specified keyword etc. is associated with an entire original and stored in the storage device. When the stored original image is retrieved,

the keyword is input, the original associated with the stored keyword is retrieved from the storage device, any corresponding original image is read, and the original image is displayed on a display, etc.

[0003]

In the data processing apparatus for filing document information generated by an application, a retrieving process is performed using a keyword added in advance corresponding to each document, or using an entire document.

[0004]

[Problem to be Solved by the Invention]

Since the conventional data processing apparatus is configured as described above, when the contents of the original image stored in the storage device are referenced, and especially when the contents are roughly grasped, there is the problem that, although an original can be grasped if a part of the contents of the original image are acquired, the process of regenerating the entire contents of the original is performed, thereby requiring a long time to confirm the contents.

[0005]

The present invention has been developed to solve the above-mentioned problem, and an objective of the present invention is to analyze a read original image, extract a result in a rectangular area of an optional

size, add different attribute data based on the form of each of the extracted rectangular areas, recognize character information by analyzing the image data in the rectangular area to which the attribute data is added, add the character information and attribute data and store them with the image data in memory resources, specify original image and attribute to be retrieved for the stored original image, and control voice-reading the character information added to the area information based on the result of determining whether or not the area depending on the specified original image and attribute has been selected from the memory resources, thereby voice-reading the character information added to the area of the original image, and allowing an operator to voice-confirm the contents of the original image in a short time only by a simple operation of specifying an operator-specific attribute and an original image without the necessity to take a long time to display the image of the entire original image.

[0006]

[Means for Solving the Problem]

The first invention according to the present invention includes: read means for reading a document image on an original and outputting image data; extraction means for analyzing the image data read by the read means, and extracting a rectangular area of an

optional size from the original image; addition means for adding attribute data different based on a form of each rectangular area extracted by the extraction means; recognition means for analyzing image data in each rectangular area to which attribute data is added by the addition means, and recognizing character information; and storage means for adding the character information recognized by the recognition means and the attribute data to each area, and storing resultant data with the image data in memory resources.

[0007]

The second invention includes: output means for voice-outputting character information; first specification means for specifying an attribute of an area to be retrieved for image data stored in the memory resources; first determination means for determining whether or not an area based on the attribute specified by the first specification means is selected from the memory resources; and first control means for controlling voice-reading by the output means based on a determination result by the first determination means.

[0008]

The third invention includes: first display control means for displaying the rectangular area extracted by the extraction means and the character information recognized by the recognition means on a

display unit; instruction means for issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and edition means for correcting the rectangular area extracted by the extraction means or the character information recognized by the recognition means in response to the correct instruction from the instruction means.

[0009]

The fourth invention includes: output means for voice-outputting character information; second display control means for displaying on the display unit a list of original images stored in the memory resources and a list of attribute items to be retrieved; second specification means for specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; second determination means for determining whether or not an area based on the attribute specified by the second specification means is selected from the memory resources; and second control means of controlling voice-reading by the output means based on a determination result by the second determination means.

[0010]

The fifth invention includes: a reading step of reading a document image on an original and outputting image data; an extracting step of analyzing the image data read in the reading step, and extracting a

rectangular area of an optional size from the original image; an adding step of adding attribute data different based on a form of each rectangular area extracted in the extracting step; a recognizing step of analyzing image data in each rectangular area to which attribute data is added in the adding step, and recognizing character information; and a storing step of adding the character information recognized in the recognizing step and the attribute data to each area, and storing resultant data with the image data in memory resources.

[0011]

The sixth invention is a storage medium storing a computer-readable program including: a reading step of reading a document image on an original and outputting image data; an extracting step of analyzing the image data read in the reading step, and extracting a rectangular area of an optional size from the original image; an adding step of adding attribute data different based on a form of each rectangular area extracted in the extracting step; a recognizing step of analyzing image data in each rectangular area to which attribute data is added in the adding step, and recognizing character information; and a storing step of adding the character information recognized in the recognizing step and the attribute data to each area,

and storing resultant data with the image data in memory resources.

[0012]

The seventh invention includes: a first specifying step of specifying an attribute of an area to be retrieved for the image data stored in the storing step; a first determining step of determining whether or not an area based on the attribute specified in the first specifying step is selected from the memory resources; and an outputting step of voice-outputting character information based on the determination result in the first determining step.

[0013]

The eighth invention is a storage medium storing a computer-readable program including: a first specifying step of specifying an attribute of an area to be retrieved for the image data stored in the storing step; a first determining step of determining whether or not an area based on the attribute specified in the first specifying step is selected from the memory resources; and an outputting step of voice-outputting character information based on the determination result in the first determining step.

[0014]

The ninth invention includes: a first display step of displaying the rectangular area extracted in the extracting step and the character information

recognized in the recognizing step on a display unit; an instructing step of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and an editing step of correcting the rectangular area extracted in the extracting step or the character information recognized in the recognizing step in response to the correct instruction in the instructing step.

[0015]

The tenth invention is a storage medium storing a computer-readable program including: a first display step of displaying the rectangular area extracted in the extracting step and the character information recognized in the recognizing step on a display unit; an instructing step of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and an editing step of correcting the rectangular area extracted in the extracting means or the character information recognized in the recognizing step in response to the correct instruction in the instructing step.

[0016]

The eleventh invention includes: a second display step of displaying on the display unit a list of original images stored in the storing step and a list of attribute items to be retrieved; a second specifying step of specifying a desired original image and



attribute item from the original images and attribute items displayed on the display unit; a second determining step of determining whether or not an area based on the attribute specified in the second specifying step is selected from the memory resources; and an output step of voice-outputting character information based on a determination result in the second determining step.

[0017]

The twelfth invention is a storage medium storing a computer-readable program including: a second display step of displaying on the display unit a list of original images stored in the storing step and a list of attribute items to be retrieved; a second specifying step of specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; a second determining step of determining whether or not an area based on the attribute specified in the second specifying step is selected from the memory resources; and an output step of voice-outputting character information based on a determination result in the second determining step.

[0018]

[Embodiments of the Invention]

[First Embodiment]

Figure 1 is a block diagram for explanation of the configuration of the data processing apparatus according to the first exemplary embodiment of the present invention.

[0019]

In Figure 1, a CPU 101 performs control of the entire data processing apparatus according to the program stored in ROM 102. The ROM 102 stores a program based on the flowchart shown in Figures 3, 4, 8, and 10 described later and executed by the CPU 101, a control program of the present apparatus, a program for processing (including a character recognizing process, an area extracting process, etc.) an image containing graphics, images, characters, tables, etc.

[0020]

RAM 103 functions as main memory, a work area, etc. of the CPU 101, and stores a document, an image, etc. An external storage device 104 is configured by storage media such as a magnetic disk (a hard disk, a floppy disk, an MO disk). A display 105 displays the process contents etc. of a data processing apparatus. A keyboard 106 and a pointing device 107 are operated to input data and issue an icon instruction to a data processing apparatus.

[0021]

An image scanner 108 (or a scanner simply) optically reads an image of an original etc., and

outputs image data to the RAM 103 or the external storage device 104. A speech synthesis unit 109 receives a character code and outputs voice information to a speaker 110. The speaker 110 voice-outputs the voice information acquired from the speech synthesis unit 109, etc.

[0022]

With the configuration of the data processing apparatus, the CPU 101 can analyze an original read as an image by the image scanner 108 etc. based on the character recognition program stored in the ROM 102, and recognize a sentence described in the original as a character code.

[0023]

Described below with reference to Figure 2 is the process performed by the CPU 101 of reading an original through the image scanner 108, and extracting an area from the read original and the display status of the original displayed on the display 105 when the area is extracted.

[0024]

Figure 2 shows the first original display window displayed on the display 105 shown in Figure 1.

[0025]

In Figure 2, an original display window W1 displays an original P1 including mixed various sentences and images in different size of characters.

Areas 1 to 8 are extracted from the original P1. Each of the areas 1 to 8 is enclosed by frame lines L1 to L8, and displayed such that each area can be distinguished from other areas.

[0026]

Described below is an example of the extracting process in each of the areas 1 to 8 in the original P1 shown in Figure 2.

[0027]

First, the original P1 as an input image is divided into  $M \times N$  small areas. If there is an image in at least one pixel in the divided small area, then it is determined that there is an image in the small area, and a thinned image is generated.

[0028]

Next, portions in which thinned images are connected are grouped to generate a small rectangle. There is a strong possibility that a rectangle having a large aspect ratio and shorter sides close to each other is a character line. Therefore, these rectangles are combined. There is also a strong possibility that a set of rectangles having almost equal shorter side lengths and arranged substantially at equal intervals refer to a text portion. Therefore, the set of rectangles is combined as one area. As a result, the areas 1 to 8 are extracted.

[0029]

Described next is the process of adding an attribute to an extracted area. It is determined that the area 2 is probably a text portion from the area extracting step (generating by merging a plurality of small rectangles), and an attribute of "text" is added. Similarly, the attribute of "text" is added to the areas 4 and 6.

[0030]

Next, for the remaining areas, it is determined whether or not a character is included. Since an area including a character periodically shows a small rectangle indicating the presence/absence of an image, an area including a character can be selected using the characteristic. As a result, the areas 1, 5, and 7 are selected as areas including characters.

[0031]

If the area further satisfies the determination condition that it is larger in line width than the character line of a text portion, and are arranged in a position different from the column of the text portion, then an attribute of a "header" is automatically added. If it satisfies the determination condition that it is larger in line width than the character line of the text portion, and is substantially equal in width of the area to the column of the text portion, then an attribute of a "sub-header" is automatically added. If it satisfies the determination condition that it is

larger in line width than the character line of the text portion, and is positioned at the lower end position or the upper end position of the original, then an attribute of a "page" is automatically added.

[0032]

As a result, the attribute of "header" is automatically added to the area 1, the attribute of "sub-header" is automatically added to the area 5, and the attribute of "page" is automatically added to the area 7. With the remaining areas, if the area is smaller, it is determined as "noise". Otherwise, it is recognized as a picture or a photo, and determined as "image". Since the area 3 is sufficiently large, the attribute of "image" is added, and the attribute of "noise" is added to the area 8.

[0033]

Thus, by automatically adding the above-mentioned attribute to an extracted area, it can be easily determined which attribute each area has for the original.

[0034]

The characteristic configuration of the present embodiment is described below with reference to Figure 1, etc.

[0035]

With the above-mentioned configuration including: read means (image scanner 108) for reading a document

image on an original and outputting image data;  
extraction means (CPU 101 for performing an area  
extracting process) for analyzing the image data read  
by the read means, and extracting a rectangular area of  
an optional size from the original image; addition  
means (CPU 101) for adding attribute data different  
based on a form of each rectangular area extracted by  
the extraction means; recognition means (CPU 101 for  
performing a character recognizing process) for  
analyzing image data in each rectangular area to which  
attribute data is added by the addition means, and  
recognizing character information; and storage means  
(CPU 101 for controlling the RAM 103 and the external  
storage device 104) for adding the character  
information recognized by the recognition means and the  
attribute data to each area, and storing resultant data  
with the image data in memory resources (RAM 103 and  
external storage device 104), a character information  
area that can be retrieved can be segmented using voice  
as a key from the original image to be registered and  
automatically stored with the registered information,  
the process of user inputting retrieval information  
about the registered image can be reduced, and an  
effective registered information appropriate for  
associating the registered image can be simply input.  
[0036]

Furthermore, with the configuration including:  
output means (speech synthesis unit 109 and speaker 110) for voice-outputting character information; first specification means (keyboard 106 and pointing device 107) for specifying an attribute of an area to be retrieved for image data stored in the memory resources; first determination means (CPU 101 for performing the process based on the flowchart shown in Figure 4 described later) for determining whether or not an area based on the attribute specified by the first specification means is selected from the memory resources; and first control means (CPU 101 for controlling the speech synthesis unit 109) for controlling voice-reading by the output means based on a determination result by the first determination means, the contents of the registered image can be efficiently checked by voice in a simple operation of an operator specifying an area to which an attribute is added without the necessity to take a long processing time to display an image of the entire image data.

[0037]

Figure 3 is a flowchart showing an example of the first data processing procedure for the data processing apparatus according to the present invention, and corresponds to the procedure of reading an original from the image scanner 108 shown in Figure 1, and



storing the original in the external storage device 104.

(1) to (5) indicate the respective steps.

[0038]

First, an image of the original P1 is input by an optoelectronic conversion by the image scanner 108 etc.

(1), the area of the input image data is extracted in the area extracting procedure (2), and an attribute (text, header, sub-header, page, etc.) is added for each of the divided areas 1 to 8.

[0039]

Next, a well known character recognizing process is performed on an area having an attribute including a character, the character information acquired from the recognition result is added to the area (4), and the information about the position, size, etc. of the obtained area, the attribute of the area, and the character information, if any, are stored in the external storage device 104 etc. (5), thereby terminating the process.

[0040]

By repeating the process, a plurality of originals read from the image scanner 108 are stored as a combination of area information in the external storage device 104. Afterwards, an important portion of the image information stored using as a retrieval key an attribute specified and input can be retrieved.

[0041]

Figure 4 is a flowchart showing an example of the second data processing for the data processing apparatus according to the present invention, and corresponds to the procedure of retrieving the original information stored in the external storage device 104 shown in Figure 1. (1) to (3) indicates the respective steps.

[0042]

First, using the keyboard 106 and the pointing device 107, the attribute based on which retrieval is to be performed is specified, the area having the same attribute as the specified attribute is retrieved from the original information stored in the external storage device 104, the corresponding area is selected (1), it is determined whether or not there is the selected area (2), and the process terminates if it is determined that there is no corresponding area.

[0043]

If it is determined in step (2) that there is a corresponding area, the character code of the character information about the selected area is transmitted from the external storage device 104 to the speech synthesis unit 109 and a speech synthesizing operation is performed, and voice-output is performed through the speaker 110 (3), thereby terminating the process.

[0044]

Thus, for example, if the attribute of "header" is specified in step (1), and the original P1 shown in Figure 2 is detected, then the area 1 is selected. As a result, the "Relating to Speech Synthesizing Technology" as the character information added to the area 1 shown in Figure 2 is voice-read through the speaker 110.

[0045]

The characteristic configuration of the present embodiment is described below with reference to Figure 4.

[0046]

With the above-mentioned configuration including: a reading step (step (1) shown in Figure 3) for reading a document image on an original and outputting image data; an extracting step (step (2) shown in Figure 3) of analyzing the image data read in the reading step, and extracting a rectangular area of an optional size from the original image; an adding step (step (3) shown in Figure 3) of adding attribute data different based on a form of each rectangular area extracted in the extracting step; a recognizing step (step (4) shown in Figure 3) of analyzing image data in each rectangular area to which attribute data is added in the adding step, and recognizing character information; and a storing step (steps (4) and (5) shown in Figure 3) of adding the character information recognized in the

recognizing step and the attribute data to each area, and storing resultant data with the image data in memory resources, a character information area that can be retrieved is segmented using voice as a key from an original image to be registered and automatically stored with the registered information, a user input process of the retrieval information about the registered image can be reduced, thereby simply inputting an effective retrieval information appropriate for associating the registered image.

[0047]

In addition, with the configuration including: a first specifying step (not shown in the attached drawings) of specifying an attribute of an area to be retrieved for the image data stored in the storing step; a first determining step (step (2) shown in Figure 4) of determining whether or not an area based on the attribute specified in the first specifying step is selected from the memory resources; and an outputting step (step (3) shown in Figure 4) of voice-outputting character information based on the determination result in the first determining step, the contents of the registered image can be efficiently checked by voice in a simple operation of an operator specifying an area to which an attribute is added without the necessity to take a long processing time to display an image of the entire image data.

[0048]

As described above, according to the present embodiment, a simple operation of an operator appropriately specifying a desired specific attribute after adding the attribute and storing the original allows the original information in the area stored with the attribute to be read, and the contents of the original can be simply checked without regenerating the contents of the original as an image on the display 105.

[0049]

Reading an image is not limited to the process performed using the image scanner (scanner 108), but the image can be stored in advance in the external storage device 104, etc. The area information can also be stored in the RAM 103. Additionally, an original stored in another data processing apparatus, a storage device, etc. in a remote position can be read using communication means.

[0050]

Furthermore, speech synthesis means is not limited to the speech synthesis unit 109, but voice information can be generated by the CPU 101 reading a supplied program.

[0051]

The area extracting method and the attribute adding method are not limited to the method according to the present embodiment so far as a similar effect

can be obtained, and it is obvious that the type of attribute is not limited to what is described with reference to the present embodiment.

[0052]

[Second Embodiment]

In the above-mentioned first embodiment, the data processing apparatus analyzes an original image, extracts each area, automatically adds an appropriate attribute, adds character information character-recognized to an area including a character, and stores the data in the external storage device 104, but an area, an attribute, and character information automatically obtained by the data processing apparatus can be corrected by a user. An embodiment is described below.

[0053]

The data processing apparatus according to the present embodiment has the same configuration as in the first embodiment shown in Figure 1.

[0054]

A second original display window displayed on the display 105 when a user corrects an area extracted by the data processing apparatus and a graphical user interface on the screen are described below with reference to Figure 5.

[0055]

Figure 5 is an explanatory view of an example of the second original display window displayed on the display 105 shown in Figure 1. The same components as those shown in Figure 2 are assigned the same reference numerals. The original displayed in the original display window is the same as that shown in Figure 2, but the extraction result is different.

[0056]

In Figure 5, W2 is an original display window, and the original P1 including various types of sentences in different character and image sizes is shown. 1a, 1b, 2 to 8 are areas extracted from the original P1. Each of the areas 1a, 1b, 2 to 8 are enclosed by the frame lines L1a, L1b, L1 to L8 respectively, and displayed such that they can be discriminated from other areas. The areas 1a and 1b correspond to the case in which the sentence of the header is divided and extracted into two portions of "Relating to Speech Synthesizing" and "Technology". Although not shown in Figure 5, it is assumed that the attribute and the character information corresponding to the extracted area are displayed in the window W2, and the information can be checked.

[0057]

In this original display window W2, for example, when the area 1b is selected, the area 1b is pointed and clicked using the pointing device 107, etc. The

selected area 1b is displayed in the frame line L1b displayed thicker than other areas. It is preferable that the selected area is easily recognized.

[0058]

When the size of the area 1a is changed, the boundary (frame line L1a) of the area 1a is pointed by the pointing device 107, the frame line L1a is held by the dragging operation and dragged to a desired position. When a new area is to be generated because a necessary area cannot be automatically extracted, the new area can be generated using the pointing device 107.

[0059]

For example, when the header portion extracted as divided areas 1a and 1b as shown in Figure 5, is combined as one area, the area 1b is first selected and deleted, and the "Technology" included in the area 1b is dragged by the pointing device 107 to be included in the frame line L1a of the area 1a, and the size of the frame line L1a of the area 1a is changed. Thus, the display is performed such that the frame line L1 can enclose the entire header in the changed area 1a, that is, the area 1. Thus, the area information, the attribute, etc. about the area 1 can be changed.

[0060]

As shown in Figure 5, an area can be corrected in adding, deleting, changing operations, etc. by a simple



operation instruction by a user using the pointing device 107 on the original display window W2.

[0061]

The window displayed on the display 105 when the attribute added to an area is edited is described below with reference to Figure 6.

[0062]

Figure 6 is an explanatory view of an example of an attribute correction window displayed on the display 105 shown in Figure 1.

[0063]

In Figure 6, W3 is an attribute correction window, and is a screen opened by a double-clicking operation by the pointing device 107 etc. on an attribute correction item registered in the pull-down menu etc. to correct the attribute, and is configured by a OK button B1, a cancel button B2, and an attribute correction box BOX1.

[0064]

The attribute correction box BOX1 is configured by: an item Z1 to be selected when the attribute is corrected to "text"; an item Z2 to be selected when the attribute is corrected to "header"; an item Z3 to be selected when the attribute is corrected to "sub-header"; an item Z4 to be selected when the attribute is corrected to "page"; an item Z5 to be selected when the attribute is corrected to "image"; and an item Z6

to be selected when the attribute is corrected to "noise".

[0065]

In the attribute correction window W3, the pointing device 107 is operated to select the items Z1 to Z6 in the attribute correction box BOX1, and the attribute can be corrected by clicking the OK button B1.

[0066]

For example, when the attribute of the area 1a shown in Figure 5 is "text", the area 1a is selected, and a correct attribute is selected in the attribute correction window W3, thereby the correction can be performed. The attribute correction window W3 can be displayed near the area whose attribute name is displayed, or the color of the area for each attribute is determined and displayed so that the attribute of an area can be easily checked.

[0067]

Thus, when the attribute of the area automatically added is different from the actual attribute, it can be corrected by the user selection using the pointing device 107 etc. in the attribute correction window W3.

[0068]

A window displayed on the display 105 when character information added to an area is edited is described below with reference to Figure 7.

[0069]

Figure 7 is an explanatory view of an example of a character information window displayed on the display 105 shown in Figure 1.

[0070]

In Figure 7, W4 is a character information window, and is a screen opened by a double-clicking operation by the pointing device 107, etc. on a character information correction item registered in the pull-down menu, etc. to correct character information, and is configured by a termination button B3, and a character information correction box BOX2, and character information MJ is displayed in the character information correction box BOX2. The portion of "C" in the character information MJ is displayed. The correction of the character information can be terminated by clicking the termination button B3.

[0071]

In Figure 7, the character information about the area 1 shown in Figure 5 is generated by the character recognizing function of a computer, and corresponds to the case in which the "Relating to Speech Synthesis Technology" is mistaken for "Relating to Speech Cynthesis Technology"

[0072]

When the above-mentioned mistake occurs, the character information can be corrected by rewriting the character information in the character information

correction box BOX2 by operating the keyboard 106 or the pointing device 107 in the character information window W4 when a mistake occurs.

[0073]

For example, when the area 1 includes "... Cyn ..." as shown in Figure 5, the area 1 is selected, the character information window W4 is displayed, and the "C" is corrected to "S".

[0074]

Thus, by operating the pointing device 107, etc. in the character information window W4, a user can directly input corrected or changed data when an automatically recognized character information in an area is different from actual character information.

[0075]

The characteristic configuration of the present embodiment is described below with reference to Figures 1, 5, 6, 7, etc.

[0076]

With the above-mentioned configuration including: first display control means for displaying the rectangular area extracted by the extraction means and the character information recognized by the recognition means on a display unit (display 105); instruction means (instruction by the keyboard 106 and the pointing device 107 in the original display window W2, attribute correction window W3, and character information window

W4) for issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and edition means (CPU 101 for processing data based on the flowchart shown in Figure 8 as described later) for correcting the rectangular area extracted by the extraction means or the character information recognized by the recognition means in response to the correct instruction from the instruction means, an operator can correct, delete, etc. a rectangular area, attribute data, and character information in a simple operation although there is a fault in an extracted rectangular area, its attribute data, and character information.

[0077]

Another embodiment relating to storing area information is described below with reference to the flowchart shown in Figure 8.

[0078]

Figure 8 is a flowchart of an example of the third data processing procedure in the data processing apparatus according to the present invention. (1) to (8) indicate the respective steps.

[0079]

As in the first embodiment, an image of an original is read using the scanner 108 (1), the area of the read original is extracted as described above (2), an attribute is automatically added to the extracted

area (3), and character information is added (4). At this time, the extracted area, attribute, and character information are displayed on the display 105, and the information can be checked.

[0080]

Next, the extracted area is edited in the original display window W2 shown in Figure 5 (5). Next, the attribute of each area is corrected in the attribute correction window W3 (6).

[0081]

Next, character information is edited in the character information window W4 shown in Figure 7 (7), the original is stored in the external storage device 104 (8) in the combination of the area, attribute, and character information edited and corrected in steps (5) to (7), thereby terminating the process.

[0082]

The characteristic configuration of the present embodiment is described below with reference to Figure 8, etc.

[0083]

With the above-mentioned configuration including: a first display step (not shown in the attached drawings) of displaying the rectangular area extracted in the extracting step and the character information recognized in the recognizing step on a display unit; an instructing step (not shown in the attached

drawings) of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and an editing step (steps (5) to (7)) of correcting the rectangular area extracted in the extracting step or the character information recognized in the recognizing step in response to the correct instruction in the instructing step, an operator can correct, delete, etc. a rectangular area, attribute data, and character information in a simple operation although there is a fault in an extracted rectangular area, its attribute data, and character information.

[0084]

As described above, an area relating to defective area extraction and defective character recognition, the attribute of the area, and the character in the area can be corrected, changed, and deleted to the user-desired contents by editing and correcting the area information, attribute information, and character information according to the present embodiment.

[0085]

In the present embodiment, in steps (1) to (3) shown in Figure 8, the editing and correcting operations are performed after generating area information in advance, but all or a part of the area information can be manually generated, and the order of

the steps (5) to (7) shown in Figure 8 is not limited to the order indicated in the present embodiment.

[0086]

[Third Embodiment]

In the above-mentioned second embodiment, an automatically determined area, attribute, and character information are corrected as necessary, and then stored in the external storage device 104, but it is first checked whether or not the area of each of a plurality of originals stored in the external storage device 104 has the same attribute as the specified attribute, and then plural pieces of character information of the area can be sequentially read. The embodiment is described below.

[0087]

The data processing apparatus according to the present embodiment has the same configuration as the first embodiment shown in Figure 1.

[0088]

The retrieval windows displayed on the display 105 when a plurality of originals stored in the external storage device 104 and the graphic user interface on the screen are described below with reference to Figure 9.

[0089]

Figure 9 is an explanatory view of an example of a retrieval setting window displayed on the display 105



shown in Figure 1. The same components as those shown in Figure 6 are assigned the same reference numerals.

[0090]

In Figure 9, W5 is a retrieval setting window opened when an instruction to execute retrieval is issued by the keyboard 106 or the pointing device 107, and is configured by the attribute specification portion BOX3 and the original specification portion BOX4. The attribute specification portion BOX3 displays the attribute setting items Z1 to Z6. The original specification portion BOX4 displays the originals 1 to 5 stored in the external storage device 104, etc. using the icons F1 to F5, and corresponds to the original specified with the inverse icons F2, F3, and F5 to be retrieved.

[0091]

In the retrieval setting window W5 configured as described above, an attribute to be retrieved can be set by selecting the attribute setting items Z1 to Z6 in the attribute specification portion BOX3 by operating the keyboard 106 or the pointing device 107, and an original to be retrieved can be selected by selecting the icons F1 to F5 in the original specification portion BOX4. B2 is a cancel button, and a retrieving process can be suspended by pressing the cancel button B2. B1 is an OK button, and the retrieval is started when the OK button B1 is pressed.

[0092]

The original specification portion BOX4 is not limited to the portion according to the present embodiment, but can be configured to display a list of names, display a reduced image on the leading page, and to use a part of the character information in the original to identify the original.

[0093]

Thus, an attribute and original to be retrieved can be easily set by operating the pointing device 107 etc. in the retrieval setting window W5.

[0094]

The characteristic configuration of the present embodiment is described below with reference to Figures 1, 9, etc.

[0095]

With the above-mentioned configuration including: output means (speech synthesis unit 109 and speaker 110) for voice-outputting character information; second display control means (CPU 101 for displaying the retrieval setting window W5) for displaying on the display unit a list of original images stored in the memory resources and a list of attribute items to be retrieved; second specification means (keyboard 106 and pointing device 107) for specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; second

determination means (CPU 101 for determining based on the flowchart shown in Figure 10 as described later) for determining whether or not an area based on the attribute specified by the second specification means is selected from the memory resources; and second control means (CPU 101 for controlling the speech synthesis unit 109) of controlling voice-reading by the output means based on a determination result by the second determination means, a registered image can be listed and displayed, and the contents of the registered image can be sequentially and efficiently retrieved by voice in a less operation.

[0096]

The fourth data processing procedure is described below with reference to Figure 10.

[0097]

Figure 10 is a flowchart of an example of the fourth data processing procedure for use with the data processing apparatus according to the present invention. (1) to (7) indicate the respective steps.

[0098]

First, the attribute to be retrieved is specified in the attribute specification portion BOX3 of the attribute correction window W3 (1), the original to be retrieved is specified in the original specification portion BOX4 (2), and it is determined whether or not all of the specified originals, for example, the icons

F2, F3, and F5 have all been retrieved (3). If it is determined that all originals have been retrieved, then the process terminates.

[0099]

On the other hand, if it is determined that all originals have not been retrieved, then control is passed to step (4). The steps (4) to (6) are similar to those in the first embodiment, and if there is an area of a specified attribute, the character information is read and output through the speaker 110.

[0100]

Then, it is determined whether or not an instruction from an operator to suspend the process has been detected, that is, whether or not the cancel button B2 shown in Figure 9 has been pressed (7). If it is determined that the button has not been pressed, control is returned to step (3). If it is determined that the button has been pressed, then the process terminates.

[0101]

As described above, an operation of reading unnecessary original information can be suppressed by setting in details the attributes and documents to be retrieved and suspending the retrieving process according to the present embodiment.

[0102]

In addition, according to the present embodiment, only an attribute and an original can be specified as a retrieval condition. However, character information can be added to the retrieval condition to specify the retrieval condition such that an area of a specified attribute and including a specified character string in the character information, or an area of a specified attribute or having character information including a specified character string can be retrieved.

[0103]

Otherwise, retrieved original name can be read, or an original having no selected area can be notified by voice to inform that the selected area does not exist.

[0104]

Furthermore, in the present embodiment, the suspension detection timing is the point where the process of the first detected original is completed, but it is not limited to this timing.

[0105]

The characteristic configuration of the present embodiment is described below with reference to Figure 10, etc.

[0106]

With the above-mentioned configuration including: a second display step (not shown in the attached drawings) of displaying on the display unit a list of original images stored in the storing step and a list

of attribute items to be retrieved; a second specifying step (steps (1) and (2) shown in Figure 10) of specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; a second determining step (step (5) shown in Figure 10) of determining whether or not an area based on the attribute specified in the second specifying step is selected from the memory resources; and an output step (step (6) shown in Figure 10) of voice-outputting character information based on a determination result in the second determining step, a registered image can be listed and displayed, and the contents of the registered image can be sequentially and efficiently retrieved by voice in a less operation.

[0107]

The configuration of the data processing program that can be read by the data processing apparatus according to the present invention is described below with reference to the memory map shown in Figure 11.

[0108]

Figure 11 is an explanatory view of a memory map of the storage medium storing various data processing programs that can be read by the data processing apparatus according to the present invention.

[0109]

Although not shown in the attached drawings, the information for management of a program graphic stored

in the storage medium, for example, version information, an author, etc. can be stored, and the information depending on the program reading OS, etc., for example, an icon identifying and displaying a program can be stored.

[0110]

Furthermore, data depending on various programs is also managed by the directory. In addition, when a program for installing various programs in a computer, and an installing program are compressed, a decompressing program etc. can be stored.

[0111]

The functions shown in Figures 3, 4, 8, and 10 according to the present embodiment can be performed by a host computer by a program installed externally. In this case, the present invention can be applied when an information group including a program is supplied to an output device from a storage medium such as CD-ROM, flash memory, an FD, etc., or from an external storage medium over a network.

[0112]

As described above, it is obvious that the purpose of the present invention can be attained by supplying a storage medium storing a program code of software for realizing the functions of the above-mentioned embodiments to a system or a device, and the computer (or the CPU or the MPU) of the system or the device

reading and executing the program code stored in the storage medium.

[0113]

In this case, the program code read from the storage medium realizes the new function of the present invention, and the storage medium storing the program code configured the present invention.

[0114]

A storage medium for supplying a program code can be, for example, a floppy disk, a hard disk, an optical disk, a magneto optical disk, CD-ROM, CD-R, a magnetic tape, a nonvolatile memory card, ROM, EEPROM, etc.

[0115]

It is obvious that the functions of the above-mentioned embodiments are realized not only by executing a program code read by a computer, but also by the process performed by the OS (operating system) operating on the computer and performing all or a part of the actual process according to the instruction of the program code.

[0116]

Furthermore, it is obvious that the functions of the above-mentioned embodiments are realized by the process performed by a program code read from a storage medium after the program code is written to the memory in the feature expansion board inserted into the computer and the feature expansion unit connected to



the computer, and the program code performing all or a part of the actual process by the CPU, etc. provided with the feature expansion board and the feature expansion unit.

[0117]

As described above, the first invention includes: read means for reading a document image on an original and outputting image data; extraction means for analyzing the image data read by the read means, and extracting a rectangular area of an optional size from the original image; addition means for adding attribute data different based on a form of each rectangular area extracted by the extraction means; recognition means for analyzing image data in each rectangular area to which attribute data is added by the addition means, and recognizing character information; and storage means for adding the character information recognized by the recognition means and the attribute data to each area, and storing resultant data with the image data in memory resources. Therefore, a character information area that can be retrieved can be segmented using voice as a key from the original image to be registered and automatically stored with the registered information, the process of user inputting retrieval information about the registered image can be reduced, and an effective registered information appropriate for associating the registered image can be simply input.

[0118]

The second invention includes: output means for voice-outputting character information; first specification means for specifying an attribute of an area to be retrieved for image data stored in the memory resources; first determination means for determining whether or not an area based on the attribute specified by the first specification means is selected from the memory resources; and first control means for controlling voice-reading by the output means based on a determination result by the first determination means. Therefore, the contents of the registered image can be efficiently checked by voice in a simple operation of an operator specifying an area to which an attribute is added without the necessity to take a long processing time to display an image of the entire image data.

[0119]

The third invention includes: first display control means for displaying the rectangular area extracted by the extraction means and the character information recognized by the recognition means on a display unit; instruction means for issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and edition means for correcting the rectangular area extracted by the extraction means or the character information

recognized by the recognition means in response to the correct instruction from the instruction means.

Therefore, a rectangular area, attribute data, and character information can be corrected, deleted, etc. in a simple operation although there is a fault in an extracted rectangular area, its attribute data, and character information.

[0120]

The fourth invention includes: output means for voice-outputting character information; second display control means for displaying on the display unit a list of original images stored in the memory resources and a list of attribute items to be retrieved; second specification means for specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; second determination means for determining whether or not an area based on the attribute specified by the second specification means is selected from the memory resources; and second control means of controlling voice-reading by the output means based on a determination result by the second determination means. Therefore, a registered image can be listed and displayed, and the contents of the registered image can be sequentially and efficiently retrieved by voice in a less operation.

[0121]

The fifth and sixth inventions include: a reading step of reading a document image on an original and outputting image data; an extracting step of analyzing the image data read in the reading step, and extracting a rectangular area of an optional size from the original image; an adding step of adding attribute data different based on a form of each rectangular area extracted in the extracting step; a recognizing step of analyzing image data in each rectangular area to which attribute data is added in the adding step, and recognizing character information; and a storing step of adding the character information recognized in the recognizing step and the attribute data to each area, and storing resultant data with the image data in memory resources. Therefore, a character information area that can be retrieved is segmented using voice as a key from an original image to be registered and automatically stored with the registered information, a user input process of the retrieval information about the registered image can be reduced, thereby simply inputting an effective retrieval information appropriate for associating the registered image.

[0122]

The seventh and eighth inventions include: a first specifying step of specifying an attribute of an area to be retrieved for the image data stored in the storing step; a first determining step of determining

whether or not an area based on the attribute specified in the first specifying step is selected from the memory resources; and an outputting step of voice-outputting character information based on the determination result in the first determining step. Therefore, the contents of the registered image can be efficiently checked by voice in a simple operation of an operator specifying an area to which an attribute is added without the necessity to take a long processing time to display an image of the entire image data.

[0123]

The ninth and tenth inventions include: a first display step of displaying the rectangular area extracted in the extracting step and the character information recognized in the recognizing step on a display unit; an instructing step of issuing a correct instruction for the rectangular area or the character information displayed on the display unit; and an editing step of correcting the rectangular area extracted in the extracting step or the character information recognized in the recognizing step in response to the correct instruction in the instructing step. Therefore, a rectangular area, attribute data, and character information can be corrected, deleted, etc. by an operator in a simple operation although there is a fault in an extracted rectangular area, its attribute data, and character information.

[0124]

The eleventh and twelfth inventions include: a second display step of displaying on the display unit a list of original images stored in the storing step and a list of attribute items to be retrieved; a second specifying step of specifying a desired original image and attribute item from the original images and attribute items displayed on the display unit; a second determining step of determining whether or not an area based on the attribute specified in the second specifying step is selected from the memory resources; and an output step of voice-outputting character information based on a determination result in the second determining step. Therefore, a registered image can be listed and displayed, and the contents of the registered image can be sequentially and efficiently retrieved by voice in a less operation.

[0125]

Therefore, the character information added to the area of the original image is voice-read, and an operator is allowed to voice-confirm the contents of the original image in a short time only by a simple operation of specifying an operator-specific attribute and an original image without the necessity to take a long time to display the image of the entire original image.

[Brief Description of the drawings]

[Figure 1]

Figure 1 is a block diagram of the configuration of the data processing apparatus according to the first embodiment of the present invention;.

[Figure 2]

Figure 2 shows an example of the first original display window displayed on the display shown in Figure 1;

[Figure 3]

Figure 3 is a flowchart of an example of the first data processing procedure in the data processing apparatus according to the present invention;

[Figure 4]

Figure 4 is a flowchart of an example of the second data processing procedure in the data processing apparatus according to the present invention;

[Figure 5]

Figure 5 is an explanatory view of the second original display window displayed on the display shown in Figure 1;

[Figure 6]

Figure 6 is an explanatory view of an example of an attribute correction window displayed on the display shown in Figure 1;

[Figure 7]

Figure 7 is an explanatory view of an example of a character information window displayed on the display shown in Figure 1;

[Figure 8]

Figure 8 is a flowchart of an example of the third data processing procedure in the data processing apparatus according to the present invention;

[Figure 9]

Figure 9 is an explanatory view of an example of a retrieval setting window displayed on the display shown in Figure 1;

[Figure 10]

Figure 10 is a flowchart of an example of the fourth data processing procedure in the data processing apparatus according to the present invention; and

[Figure 11]

Figure 11 is an explanatory view of a memory map of a storage medium storing various data processing programs readable in the data processing apparatus according to the present invention.

[Description of Reference Numerals]

101 CPU

102 ROM

103 RAM

104 external storage device

105 display

106 keyboard



- 107 pointing device
- 108 scanner
- 109 speech synthesis unit
- 110 speaker

Figure 1

104 EXTERNAL STORAGE DEVICE  
105 DISPLAY  
106 KEYBOARD  
107 POINTING DEVICE  
108 SCANNER  
109 SPEECH SYNTHESIS UNIT

Figure 2

1 RELATING TO SPEECH SYNTHESIS TECHNOLOGY  
2 SPEECH SYNTHESIS TECHNOLOGY REFERS TO TECHNOLOGY  
OF CONVERTING SENTENCE INTO AUDIO SIGNAL. THIS  
SENTENCE IS FOR EXPLANATION OF THE PRESENT INVENTION ...  
4 SAMPLE. TEXT IN THIS EXAMPLE.  
5 SPEECH SYNTHESIS MEANS  
6 BEFORE EXPLANATION OF SPEECH SYNTHESIS, INPUTTING  
IMAGE IS DESCRIBED

Figure 3

#1 START  
1 READ IMAGE  
2 EXTRACT AREA  
3 ADD ATTRIBUTE TO AREA  
4 ADD CHARACTER INFORMATION TO AREA OF ATTRIBUTE  
INCLUDING CHARACTER  
5 STORE AREA INFORMATION  
#2 END

Figure 4

```
#1  START
1   RETRIEVE ATTRIBUTE
2   AREA SELECTED?
3   READ CHARACTER INFORMATION ABOUT SELECTED AREA
#2  END
```

Figure 5

```
1   RELATING TO SPEECH SYNTHESIS TECHNOLOGY
2   SPEECH SYNTHESIS TECHNOLOGY REFERS TO TECHNOLOGY
OF CONVERTING SENTENCE INTO AUDIO SIGNAL. THIS
SENTENCE IS FOR EXPLANATION OF THE PRESENT INVENTION ...
4   SAMPLE. TEXT IN THIS EXAMPLE.
5   SPEECH SYNTHESIS MEANS
6   BEFORE EXPLANATION OF SPEECH SYNTHESIS, INPUTTING
IMAGE IS DESCRIBED
```

Figure 6

```
21  TEXT
22  HEADER
23  SUB-HEADER
24  PAGE
25  IMAGE
26  NOISE
B2  CANCEL
```

Figure 7

#1 RELATING TO SPEECH CYNTHESIS TECHNOLOGY  
#2 TERMINATE

Figure 8

#1 START  
1 READ IMAGE  
2 EXTRACT AREA  
3 ADD ATTRIBUTE TO AREA  
4 ADD CHARACTER INFORMATION TO AREA OF ATTRIBUTE  
INCLUDING CHARACTER  
5 EDIT AREA  
6 CORRECT ATTRIBUTE  
7 EDIT CHARACTER INFORMATION  
8 STORE AREA INFORMATION  
#2 END

Figure 9

21 TEXT  
22 HEADER  
23 SUB-HEADER  
24 PAGE  
25 IMAGE  
26 NOISE  
P1 ORIGINAL 1  
P2 ORIGINAL 2  
P3 ORIGINAL 3

P3 ORIGINAL 4  
P5 ORIGINAL 5  
B2 CANCEL

Figure 10

#1 START  
1 SPECIFY ATTRIBUTE  
2 SPECIFY ORIGINAL TO BE RETRIEVED  
3 ALL ORIGINALS RETRIEVED?  
4 SPECIFY ORIGINAL TO BE RETRIEVED  
5 AREA SELECTED?  
6 READ CHARACTER INFORMATION IN SELECTED AREA  
7 RETRIEVAL SUSPENDED?  
#2 END

Figure 11

#1 STORAGE MEDIUM SUCH AS FD/CD-ROM  
#2 DIRECTORY INFORMATION  
#3 FIRST DATA PROCESSING PROGRAM  
PROGRAM CODE GROUP CORRESPONDING TO STEPS IN  
FLOWCHART SHOWN IN Figure 3  
#4 SECOND DATA PROCESSING PROGRAM  
PROGRAM CODE GROUP CORRESPONDING TO STEPS IN  
FLOWCHART SHOWN IN Figure 4  
#5 THIRD DATA PROCESSING PROGRAM  
PROGRAM CODE GROUP CORRESPONDING TO STEPS IN  
FLOWCHART SHOWN IN Figure 8

#6 FOURTH DATA PROCESSING PROGRAM

PROGRAM CODE GROUP CORRESPONDING TO STEPS IN  
FLOWCHART SHOWN IN Figure 10

#7 MEMORY MAP OF STORAGE MEDIUM